

声道模型の動きを制御する試み*

○荒井隆行（上智大・理工）

1 はじめに

人間の声道を模擬した声道模型は、その形状を変えることによって異なる音を作ることが可能となる。例えば、古くは von Kempelen の speaking machine においては、手で声道形状を変えるような作りになっている [1]。その後、ロボット技術の延長で例えば発話ロボットなどが開発され、声道形状を機械制御によって変える仕組みなどが提案されている（例えば [2-5]）。

そのような中、我々は梅田・寺西[6]による声道模型を改造し、アクチュエータを接続したものを報告した[7]。そのシステムでは、PC から時間ステップごとの位置情報を送信し、それに従って声道形状を変化させた。しかし、その性能について検討の余地が残されたままになっていた。またその後、主に教育目的に開発された解剖模型風の声道模型など、舌を中心に形状を変えられる声道模型を開発した（例えば[8]）が、それらは手で形状を変えるものであり、機械制御には至っていなかった。

そこで本稿では、改めて PC 制御による梅田・寺西式声道模型[7]を再び稼働させながらその性能を含め再検討すると共に、解剖模型風の声道模型に動きを加える試みについても取り扱う。

2 梅田・寺西式声道模型

2.1 概略

Fig. 1 に外観を示す。もともと、梅田・寺西式声道模型では、声道形状を片側から挿入された複数の角棒を抜き差しすることによって変化させることができるようになっている。図にある声道模型は梅田規子先生から寄贈されたものだが、当初は口腔には 11 本の角棒が、鼻腔には 8 本の角棒が設置されていた。そのうち、口腔の 11 本それぞれに 11 基のアクチュエータを接続し、制御装置から位置をコントロールするように改造した[7]。

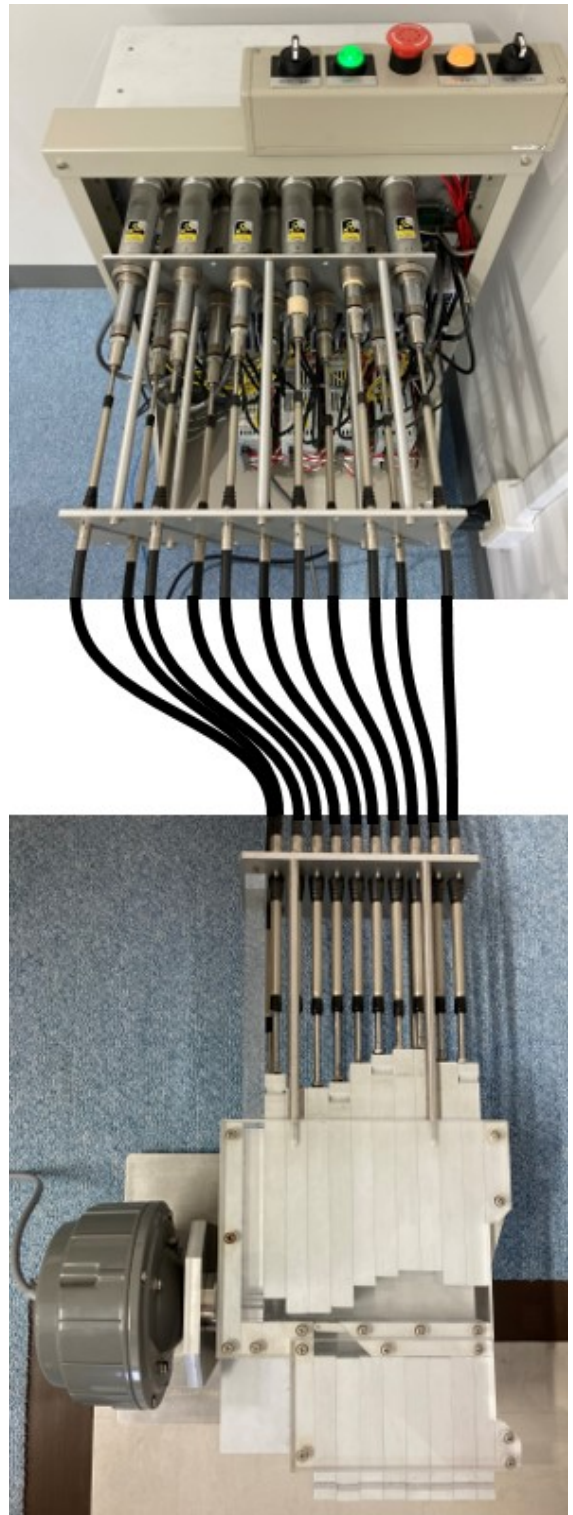


Fig. 1: PC-controlled Umeda-Teranishi vocal-tract model [7].

* Few attempts on controlling the movements of vocal-tract models, by ARAI, Takayuki, (Sophia University).

制御部については、マシンコントローラ (Yaskawa, MP2200S) を PC 上のソフトウェア (Yaskawa, MPE720 Ver. 6) から操作する仕様となっている。また、アクチュエータは IAI 社製 (RCA-RA4C-I-30-12-50-A1-PHA) で、コントローラ (ACON-C-30IHA-ML-0-0) と組みになっている。

2.2 その後の改良

もともと、Fig. 1 にある梅田・寺西式声道模型には、鼻咽腔結合をコントロールする弁が付いていた。それをアクチュエータでコントロールすることができるようにするため、喉頭側の角棒は動かない仕様に変更し、そのために空いたアクチュエータを鼻咽腔結合用の弁の開閉に用いることとした。

また制御系において、PC の OS を Windows 11 にするのに伴い、MPE720 も Ver. 7 にバージョンアップした。

2.3 動作テスト

動作テストに際し、音源として PC から喉頭原音を模擬した信号を再生し、ホーンスピーカのドライバユニット (TOA, TU-750) から出力した。ドライバユニットのネック部には中央に小さい穴の開いた円柱状の詰め物を挿入した。

動作テストとして、まず日本語 5 母音を繰り返し生成するという動作テストを行った。その結果、5 母音が比較的明瞭に繰り返し出力されるのを確認した。

次に、どの程度の子音が出力できるかを評価するため、まずは口唇側の角棒を最速で開閉するテストを行った。その際、その他のアクチュエータは母音[a]のままとした。出力音については、[wa]に近い音となっていた。アクチュエータの位置を格納するテーブルにおいて、100 行の動作を数回繰り返しその周期から逆算したところ、1 行ごとの位置命令につき平均で約 31.4 ms を要していた。

さらに、母音[i]から最速で母音[a]の位置にアクチュエータを動作させた。その結果、出力音については[ja]に近い音となった。

最後に、鼻咽腔結合をさせながら[wawa]と同じ調音動作を再現した。その際、母音区間で鼻咽腔閉鎖をさせる場合とさせない場合の両方を試した。その結果、[mama]と

聞こえるような出力音は得られなかった。これは、アクチュエータの速度の限界以外に、鼻咽腔閉鎖時にも出力音が常に鼻音化したような響きになってしまっていることが原因と推察された。

3 解剖模型風声道模型

3.1 概略

文献[8]で報告した解剖模型風声道模型は、人間の声道により近く途中で直角に曲がった声道形状をしている。また、舌を軟らかい素材で実現することにより、声道形状をある程度、変えることが可能である。その他、位置は固定されているものの顎や歯、口唇などがあるため、各器官と声道の関係が目でも確認できるのが特徴である。そこに3つのモータとアームとなる部品等を組み合わせ、舌の形状を変えることを試みた。Fig. 2 には、その断面を示す。また、Fig. 3 には3つのモータの動きを組み合わせることで作り出される、3つの想定される舌形状を示す (ピンクは舌+ α の部品)。

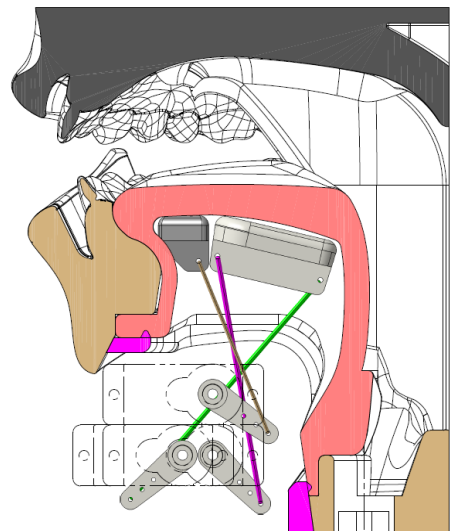


Fig. 2: Midsagittal cross-sectional view of the PC-controlled anatomical model.

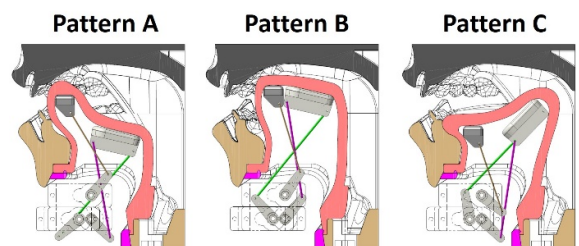


Fig. 3: Three major tongue configuration patterns.

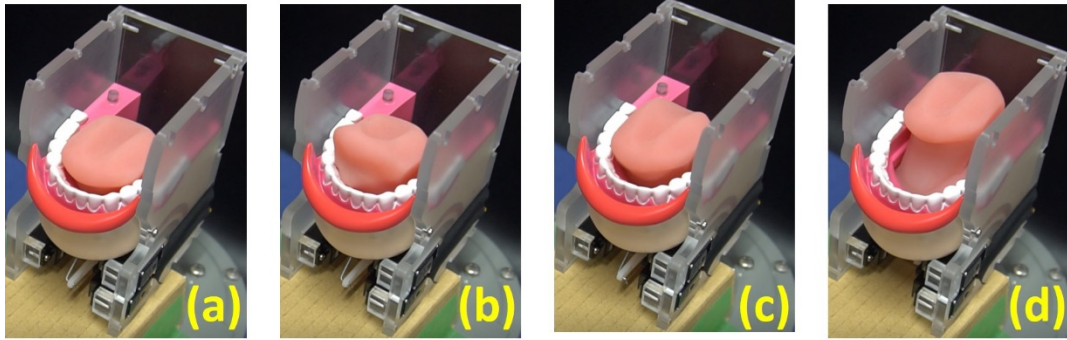


Fig. 4: Tongue configurations: (a) in the resting position, (b) Pattern A, (c) Pattern B, and (d) Pattern C.

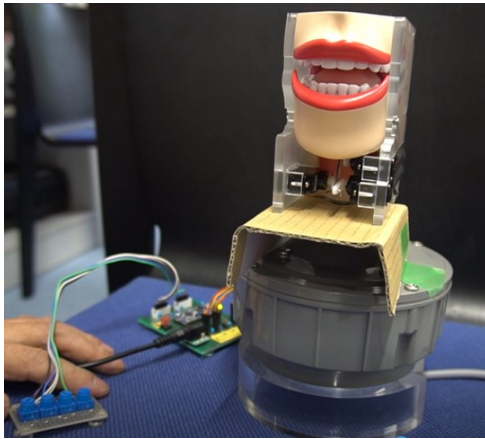


Fig. 5: Experimental settings for speech production.

Fig. 2 および Fig. 3 を見ると分かるように、舌の下側には3つのモータとそれぞれに接続された3つのアームがあり、各アームの角度によって異なる舌形状が実現されている。Fig. 3において、Pattern Aは舌尖が歯茎に向かう形状、Pattern Bは舌体が口蓋に向かう形状、Pattern Cは舌背の奥が口蓋垂の方向へ向かう形状となっている。

3.2 動作テスト

前節の解剖模型風声道模型について、実際に動作テストを行った。Fig. 4は、Fig. 2のneutralな舌形状に加えてFig. 3の3つのパターンを実際に動かした様子を示す。なお、この図では口蓋を含む上顎の部品は取り外してある。Fig. 4(b)ではPattern Aが比較的实现できている一方、Fig. 4(c)ではPattern Bが途中までしか実現できていなかった。つまり、母音[i]ほどの舌の高さは実現出来ておらず、母音[e]程度に留まっていた。また、Pattern Cは母音[u]を想定した形状だが、Fig. 4(d)からはやはり動きが途中ま

で、舌の動きが完全には目標位置に到達していないことが分かった。

Fig. 5に、音声生成のための実験の様子を示す。模型の下にはホーンスピーカのドライバユニット (TOA, TU-750) を設置し、喉頭原音となる音源をPCから再生した。モータはPCからの指令で制御可能であると同時に、基板上的可変抵抗を使ってマイコンボードを介しても制御が可能となっている。生成された音を確認したが、母音の明瞭度はあまり高くはなかった。

4 おわりに

本稿では、声道模型に対しコンピュータ制御によって声道形状を変えることを改めて試みた。1つめとして、PC制御を可能にした梅田・寺西式声道模型を再び検証した。その結果、日本語5母音と限られた子音が生成されたが、同時に限界も確認された。2つめとして、解剖模型風声道模型に対しモータを組み合わせることによって舌形状をコントロールすることを試みた。後者については、舌の動きをある程度は変えることが出来た一方、音としては明瞭な母音を生成し分けるには至らなかった。これらの試みから、ある程度のスピードを考慮したアクチュエータを選別し、さらに舌を含む調音器官を動かす機構について工夫を重ねることによって改良が進められるであろうという感触は得られた。そこで今後は、その改良を重ねる予定である。

謝辞

内容の一部は、JSPS 科研費 21K02889 ならびに上智大学重点領域研究の助成を得た。

参考文献

- [1] Dudley, H. and Tarnoczy, T. H., “The speaking machine of Wolfgang von Kempelen,” *J. Acoust. Soc. Am.*, 22(2), 151–166, 1950.
 - [2] Mochida, T., Honda, M., Hayashi, K., Kuwae, T., Tanahashi, K., Nishikawa, K. and Takanishi, A., “Control system for talking robot to replicate articulatory movement of natural speech,” *Proc. INTERSPEECH*, 1533–1536, 2002.
 - [3] Hofe, R. and Moore, R. K., “Towards an investigation of speech energetics using ‘AnTon’: an animatronic model of a human tongue and vocal tract,” *Connection Science*, 20(4), 319–336, 2008.
 - [4] Fukui, K., Kusano, T., Mukaeda, Y., Suzuki, Y., Takanishi, A. and Honda, M., “Speech robot mimicking human articulatory motion,” *Proc. INTERSPEECH*, 1021–1024, 2010.
 - [5] Sawada, H. and Hashimoto, S., “Mechanical model of human vocal system and its control with auditory feedback,” *JSME International Journal, Series C*, 43(3), 645–652, 2000.
 - [6] Umeda, N. and Teranishi, R. “Phonemic feature and vocal feature: Synthesis of speech sounds, using an acoustic model of vocal tract,” *J. Acoust. Soc. Jpn.*, 22(4), 195–120, 1966.
 - [7] Arai, T., “Mechanical vocal-tract models for speech dynamics,” *Proc. INTERSPEECH*, 1025–1028, 2010.
 - [8] Arai, T., “Flexible tongue housed in a static model of the vocal tract with jaws, lips and teeth,” *Proc. INTERSPEECH*, 171–172, 2018.
-